This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2014.2332001, IEEE Journal of Biomedical and Health Informatics

Electrocardiogram Classification using Reservoir Computing with Logistic Regression

Miguel Escalona-Morán, *Member, OSA,* Miguel C. Soriano, *Member, IEEE,* Ingo Fischer, and Claudio R. Mirasso

***Abstract*—An adapted state-of-the-art method of processing information known as Reservoir Computing is used to show its utility on the open and time consuming problem of heartbeat classification. The MIT-BIH arrhythmia database is used following the guidelines of the Association for the Advancement of Medical Instrumentation (AAMI). Our approach requires a computationally inexpensive pre-processing of the electrocardiographic signal leading to a fast algorithm and approaching a real-time classification solution. Our multi-class classification results indicate an average specificity of 97.75% with an average accuracy of 98.43%. Sensitivity and positive predicted value show an average of 84.83% and 88.75%, respectively, what makes our approach significant for its use in a clinical context.**

***Index Terms*—Reservoir Computing, ECG classification, delay system, logistic regression**

I. Introduction

FOR more than 50 years, electrocardiograms (ECG) have been a powerful and irreplaceable tool in the exploration and diagnostic of cardiovascular diseases. Its acquisition requires only simple and low-cost devices with a minimum impact on the patient. This has made ECG a preferred tool used in clinical environments and outside by the use of portable devices such as the Holter monitor. This popularity results in a large amount of data to be analyzed by physicians in a time consuming process. The ECG is an essential diagnostic tool for common pathologies such as myocardial ischemia [1], [2], arrhythmia and other rare pathologies as cardiac muscular dystrophy or Brugada syndrome [3].

Several computational algorithms have been proposed to automate the process of ECG classification. The usual ECG classification solution includes a multistep procedure. A first step of detecting the heartbeat, usually done by a QRS detector algorithm, such as the Pan Tompkins method [4]. A second step requires features extraction, transforming the raw signal to meaningful quantities, such as the ECG morphology [5], the duration of different intervals on the ECG wave [6], the separation of landmarks on the waveform among several heartbeats [5], [6], coefficients of transformation of the original signal, etc. Most of the current algorithms rely on a good feature extraction to achieve good performance. However, this step can be time consuming due to the different waveforms and the fact that pathological beats not always clearly show the landmarks needed for the extraction of features. The last step of the processing represents the classification itself. This step consists of distinguishing the different types of beats to be classified. Many authors have already studied the heartbeat classification problem using several different techniques, such as self-organizing networks (SON) [7], self-organizing maps with learning vector quantization (SOM-LVQ) [8], linear discriminants (LD) [9], [10], signal modeling (SM) [10], support vector machine (SVM) [11], [12], discrete wavelet transformation (DWT) [13], Bayesian artificial neural networks (BANN) [14], local fractal dimension [15] and delay differential equations (DDE) [16], obtaining different performance measures. Comparing results is difficult though, because of the different measures that were used, as well as the different partitions of the available data into training, testing and validation subsets. In 1987, the Association for the Advancement of Medical Instrumentation (AAMI) published a guideline for grouping heartbeat types into classes [17] and for evaluating the performance of algorithms [18]. Unfortunately, very few investigators have utilized these standards what makes the direct comparison of results and the identification of pros and cons of the different methods difficult.

In this work, we present an approach that requires a computationally inexpensive pre-processing of the electrocardiogram, leading to a fast algorithm and heading to a real-time heartbeat classification. To that end, we employ a machine learning paradigm named Reservoir Computing (RC) [19]–[25]. RC mimics brain neural networks by processing information that generates patterns of transient neural activity as a response to a sensory signal [26] and is composed of layers for processing the information. It has been used for classification tasks, time series prediction and modeling [25], [27]–[32]. In particular, it has been used for the analysis of biological signals to detect epileptic seizures [33].

We use a kind of reservoir consisting of a nonlinear dynamical element subject to delayed feedback. Nonlinear systems with delayed feedback and/or coupling, also known as delay systems, arise in a variety of real live contexts [34]. These systems can exhibit a wide range of dynamics ranging from stable operation to periodic oscillations and deterministic chaos [35]. One of the simplest delay system is a single nonlinear node influenced by its own dynamics after a certain delay time τ. In this article we utilize this simple delay system for processing information within the framework of Reservoir Computing since it allows to process information sequencially and can be easily build in hardware [36].

At least three layers are needed, an input layer to feed

Manuscript received October 19, 2013; revised December 27, 2013.

The authors are with Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain (e-mail: miguelangel@ifisc.uib-csic.es; miguel@ifisc.uib-csic.es; ingo@ifisc.uib-csic.es; claudio@ifisc.uib-csic.es).

This work was supported by the grant FIS2012-30634 (Intense@cosyp) from MINECO (Spain) and FEDER and Grups Competitius, Comunitat Autonoma de les Illes Balears, Spain.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2014.2332001, IEEE Journal of Biomedical and Health Informatics

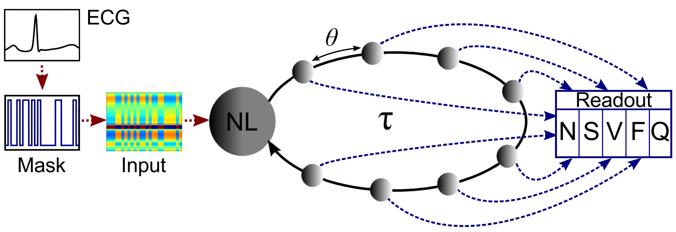

Fig. 1. Schematic view of reservoir computing based on a single nonlinear node with delay. Virtual nodes are defined as different temporal positions in the delay line. The dashed lines indicate the implicit connection between the virtual nodes and the readout which contains the 5 different classes to be classified.

the information, the reservoir layer that transforms the input through the nonlinearity and an output layer collecting the results of the processing and adapting the classifier through the learning process. The typical configuration of the reservoir is a set of nonlinear nodes. The response of these nodes to the input is read out with certain weights. The aim of the learning process is to compute the weights. In our case, we used a ring configuration (see Fig. 1) with only a single active nonlinear node and the other nodes, called *virtual nodes*, are delayed versions of the output of the nonlinearity. We come back to this point in the next section.

The manuscript is organized as follows. Section II describes the methodology of our approach. Then, the database is described in section III. Section IV contains the results of this study. At first, parameters are determined to obtain the smallest error. Then, typical machine learning evaluation measures are computed, and a comparison of the presented approach with the results of other studies is shown. Conclusions are discussed in section V.

II. METHODOLOGY AND MODEL

The reservoir is implemented using a nonlinear node subject to delayed feedback following the concept as introduced in [25]. This approach, which is easy to implement in hardware, resembles a recurrent network with ring topology [37]. As illustrated in Fig. 1, the input information is only fed into the nonlinear node. However, the transient dynamics of the system is readout at different temporal positions in the delay line, emulating temporally the existence of several processing nodes, the virtual nodes. For computation, the transient responses at predefined temporal positions are combined in the readout. All computational steps can be summarized as follows:

1) Each point of the signal to be processed is sampled and held during one delay time τ and multiplied by a binary random mask –consisting of ones and minus ones– resulting in an input matrix, I, of dimensions $M \times N$, where M is the number of sampled points in the signal to be processed (the ECG signal) and N is the number of virtual nodes in the delay line, see Fig. 1.
2) The input matrix I is then fed into the nonlinearity in a serial manner, creating a pattern of transient activity in the delay line. The response to every row of I fills the delay vector of length τ.
3) When all elements of matrix I are injected, a *state matrix* S, which contains the transient responses to the input signal at the virtual nodes, is constructed.
4) This results in one state matrix per heartbeat. These state matrices are used for constructing the classifiers through a learning process.

The learning process is a regression that allows to find the best fitting and most parsimonious, efficiently interpretable, model to describe the relationship between a dependent variable and a set of independent variables [38]. For the readout process we employ logistic regression. This is in constrast to standard Reservoir Computing which uses linear regression methods. The logistic regression (LR) [39] is a widely used learning technique in biostatistical applications in which binary responses occur quite frequently, in questions such as a condition is present or absent. LR is specified in terms of logit transformations, defined as

$$logit(P) = \ln(odds) = \ln\left(\frac{P}{1-P}\right) \tag{1}$$

where the odds represent the ratio of the probability P that an event will occur to the probability that the same event will not occur. In the logistic regression, the aim is to linearly relate the logit function with the set of states matrices containing the transient responses of the ECGs. From now on, we call D to the set of states matrices S. Then, we need to find the values of parameters a and b that satisfy

$$logit(P) = a + bD. \tag{2}$$

Consequently, results can be directly interpreted as the probability of a condition to be true or false through the following equation

$$P = \frac{e^{a+bD}}{1+e^{a+bD}}. \tag{3}$$

Note that logit functions are linearly related to the data D, but the probabilities are nonlinearly related to it. This is an advantage since in classical linear models it is usually assumed that the outcomes are independent and normally distributed with equal variance. These assumptions are often inadequate in medical applications. Finally, through LR, we are able to compute the weights of the output layer. There is one weight for each component of the S matrix. These weights are regularized using the so called L1 regularization [39] that approximate the less relevant weights to zero. In this way, the results of the LR could be interpreted as the most relevant parts of the ECG morphology that contributes to the classification.

The nonlinearity used in our system is adapted from the Mackey-Glass oscillator [40]. The model contains a delayed feedback term and has been extended to include an external input $I(t)$. Appeltant et al. [25] presented the extended model as:

$$\dot{X}(t) = -X(t) + \frac{\eta \cdot [X(t-\tau) + \gamma \cdot I(t)]}{1 + [X(t-\tau) + \gamma \cdot I(t)]^p}, \tag{4}$$

with X describing the temporal evolution of the dynamics of the nonlinear function, $\dot{X}$ being its derivative with respect

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2014.2332001, IEEE Journal of Biomedical and Health Informatics

TABLE I
TYPES OF HEARTBEATS IN MIT-BIH ARRHYTHMIA DATABASE AND THE CLASS THEY BELONG

Normal	Pathological
normal beat	atrial premature beat
left bundle branch block	aberrated atrial premature beat
right bundle branch block	nodal premature beat
	supraventricular premature beat
	atrial escape beat
	nodal escape beat
	premature ventricular contraction
	ventricular escape beat
	fusion of ventricular and normal beat
	paced beat
	fusion of paced and normal beat
	unclassified beat

to a dimensionless time t, and τ denoting the normalized delay in the feedback loop. Parameters η and γ represent feedback strength and input scaling, respectively. Without an external input ($\gamma = 0$) the system is chosen to operate in a stable fix point. However, under external inputs the system can exhibit complex dynamics. In particular, we are interested in a dynamical regime that produces consistent transient responses. The exponent p can be used to tune the nonlinearity. Although we have chosen the Mackey-Glass nonlinearity, it is expected that other nonlinear functions perform similarly well. For instance, a semiconductor laser has been used to perform similar tasks [27], [36].

As mentioned before, the virtual nodes depicted in Fig. 1 correspond to different temporal positions in the delay loop τ. In fact, the inverse of the sampling rate of the signal defines τ and the inverse rate of the binary random mask defines the temporal spacing between the virtual nodes. This temporal spacing, termed θ, is chosen such that the response of the nonlinear system remains in a transient state. We have empirically found that typically the optimum computational results are obtained for $\theta \sim 0.2$. In turn, the spacing between the virtual nodes, θ, and the number of virtual nodes, N, define the total delay in the feedback loop, $\tau = N\theta$.

III. DATA BASE DESCRIPTION

We use the MIT-BIH Arrhythmia Database [41] available at Physionet which contains 48 ambulatory ECG recordings of half hour each, obtained from 47 subjects studied by the Massachusetts Institute of Technology (MIT) and the Beth Israel Hospital (BIH) Arrhythmia Laboratory between 1975 and 1979. Twenty-three recordings were selected at random from the whole BIH database including a mix population of inpatients (about 60%) and outpatients (about 40%). The remaining 25 recordings were selected to include less common but clinically significant arrhythmias that would not be well-represented in a small random sample. The recordings were digitized at 360 samples per second with 11-bit resolution over a $10mV$ range. Two or more cardiologists independently annotated every heart beat in each record.

We used ECG lead known as modified limb lead II (MLII) which is present in 45 recordings. The database contains 15 different heartbeat types. In table I a list of the different heartbeat types in a two-class format is shown. Paced heartbeats were removed from the following analysis in agreement with AAMI guidelines.

Following the partition introduced by De Chazal, et al. [9] the database was divided into two datasets containing approximately the same amount of heartbeats and the mixture of routine and complex arrhythmia recordings. The first dataset (DS1) was used to train the classifier, while the second dataset (DS2) was used for testing. Table II shows the records included in each dataset.

To perform heartbeat classification, the ECGs were divided into heartbeats using a fixed-length window of 170 samples, this is $M = 170$ on section II, around the R-peak. This particular point of the ECG is annotated in the database. However, when using different databases, our approach needs to detect the approximate location of the R-peak in order to lock the time window. The window was positioned around the maximum peak of the QRS complex to extract the waveform. Seventy samples before the R-peak were extracted to include P waves, and 100 samples after the R-peak were also included to have information about the T wave and the duration of the heartbeat. Figure 2 shows the mean of the heartbeats of a normal subject. We marked the most relevant waves and points in a typical ECG waveform: the P wave –due to the depolarization of cardiac cells throughout the atria–, the QRS complex –representing ventricular depolarization–, and the T wave –showing ventricular repolarization.

IV. RESULTS

As stated before, logistic regression with L1 regularization is used in this study. We set the strength of the regularization to 1 and the tolerance for the convergence of the algorithm that finds the minimum of parameters a and b (eq. 2) to 10^{-6}.

First, we evaluated the number of virtual nodes needed for a proper heartbeat classification using DS1. To do so, we computed the error rate (ER), i. e. the number of misclassified heartbeats over the total number of heartbeats, for different number of virtual nodes. The exponent p of the Mackey-Glass equations (eq. 4) was choosen to be $p = 7$, since a high nonlinearity is beneficial for separating classes while not much memory is required [25]. Figure 3 shows the ER as a function of the number of virtual nodes. We computed two error rates: the training (dashed line) and the testing (continuos line) ER. The former represents the testing misclassification over the

TABLE II
MIT-BIH ARRYTHMIA DATABASE RECORDING NAMES USED FOR TRAINING AND TESTING DATASETS

Dataset	MIT-BIH Arrhythmia record names
DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230.
DS2	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234.

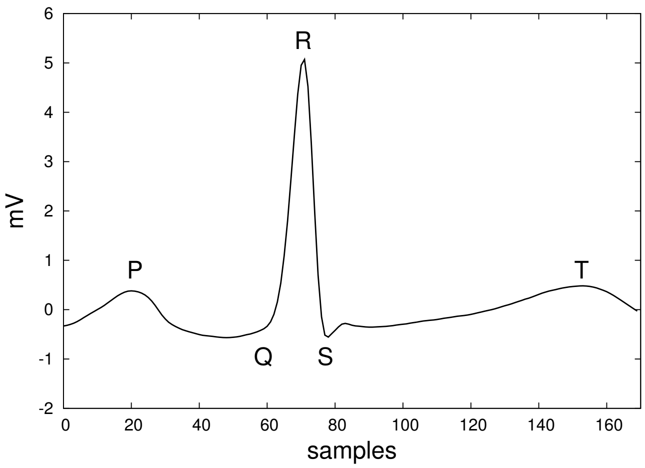

Fig. 2. ECG trace of the mean of the heartbeats for a normal subject. P wave represents the atria depolarization, QRS complex takes place at ventricular depolarization and T wave shows ventricular repolarization.

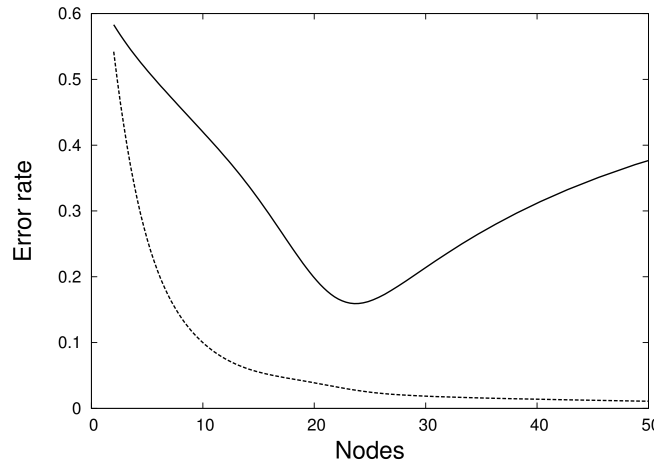

Fig. 3. Model complexity. Dependence of the training (dashed line) and testing (continouos line) error rates as a function of the number of virtual nodes in the reservoir. For less than 25 nodes, the system suffers a high bias problem. For more than 25 nodes, the system has a high variance problem. Due to this, a total number of 25 nodes is found to be the optimal.

training set (DS1) which is the same set used for training. We apply a patient oriented cross-validation where we train with all patients leaving one out. The testing error is the ER over the patient that was taking out during the training phase. For a number of nodes less than 25 the classifier suffers a high bias problem, meaning that the model is too simple to fit the data. This scenario is recognized because both ERs, training and testing, are high. In contrast, for a number of nodes greater than 25 the classifier suffers a high variance problem noticeable by the low training ER and the high testing ER. High variance problem means that the reservoir is overfitting the data and it is not capable to generalize. For 25 virtual nodes, we found a trade off between these two problems indicating the optimal number of nodes. The big difference between the training and testing ER at this point indicates a high variability of ECG morphology that could be reduced by increasing the number of samples. In what follows, a total number of 25 virtual nodes are used for the reservoir.

Second, we explored the dependence of the testing ER with the Mackey-Glass system's parameters η and γ using DS1. This is shown in Figure 4 in a logarithmic gray scale. In this figure it can be noted that a good performance of the reservoir is obtained, for a high value of the feedback strength η and a high value of the input scaling. Parameters $\eta = 0.8$ and $\gamma = 0.5$ were used for the computations, resulting in an error of $\sim 10\%$. The same test was made optimizing sensitivity, specificity, positive predicted value ($+P$) and accuracy, obtaining similar optimal values for η and γ parameters. Those results are not shown in this article.

To Compare the performance with other published results is a difficult task. One of the difficulties is the definition of classes. Some studies considered each heartbeat type as a different class (see table I). Others grouped hearbeat types to reduced the amount of classes, for instance, as recommended by the AMMI. Another problem is related to the measures of performance. These measures have advantages and disadvantages, see [42] for a review. Here, we computed the most common ones for comparison purposes.

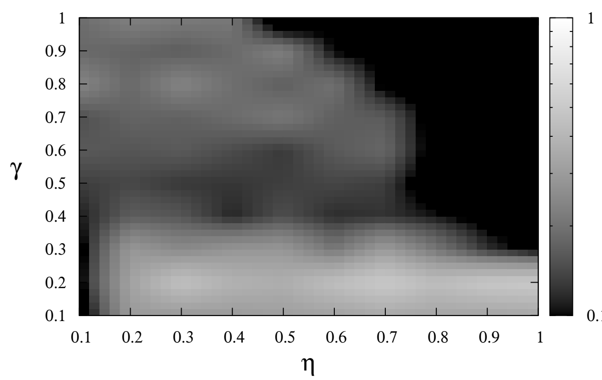

Fig. 4. Test Error Rate as a function of parameters η and γ. The exponent p was chosen to be $p = 7$. We choose as optimal parameters to run the reservoir $\eta = 0.8$ and $\gamma = 0.5$

In our approach, we use the AAMI guidelines which prompt to combine the heartbeat types (see table I) into five classes, namely, N, S, V, F and Q. Classes can contain more than one heartbeat type. Class N contains normal and bundle branch block beats, class S is the supraventricular ectopic (SVEB) class including atrial, aberrated atrial, nodal and supraventricular premature beats, as well as atrial and nodal escape beats. A ventricular ectopic (VEB) class is known as class V with premature ventricular contraction and ventricular escape beats. Class F represents fusion of ventricular and normal beats. The last class, class Q, contains the remaining beat types in the database. For performance reasons, the AAMI recommends to compute some measures that focus on the ability of the algorithm to classify VEB from non-VEB, and SVEB from non-SVEB. We computed these measures using DS2. Our results, shown in Table III, are characterized by a high accuracy ($> 97\%$), a low false positive rate ($< 2\%$). The

TABLE III
PERFORMANCE MEASURES (IN %) RECOMMENDED BY THE AAMI FOR THE EVALUATION OF CLASSIFIERS. MEASURES ARE MADE USING DS2.

Rec	number of beats					SVEB				VEB			
	N	**S**	**V**	**F**	**Q**	*Acc* (%)	*Se* (%)	*+P* (%)	*FPR* (%)	*Acc* (%)	*Se* (%)	*+P* (%)	*FPR* (%)
100	2239	33	1	0	0	99.6	93.9	79.5	0.4	100.0	100.0	100.0	0.0
103	2082	2	0	0	0	99.6	50.0	12.5	0.3	100.0	–	–	0.0
105	2526	0	41	0	5	97.2	–	0.0	2.8	99.9	95.1	100.0	0.0
111	2123	0	1	0	0	99.7	–	0.0	0.3	100.0	100.0	100.0	0.0
113	1789	6	0	0	0	97.1	83.3	8.9	2.8	100.0	–	–	0.0
117	1534	1	0	0	0	97.3	100.0	2.4	2.7	100.0	–	–	0.0
121	1861	1	1	0	0	98.5	100.0	3.6	1.5	100.0	100.0	100.0	0.0
123	1515	0	3	0	0	99.6	–	0.0	0.4	99.9	66.7	100.0	0.0
200	1743	30	826	2	0	99.3	74.4	20.9	0.6	99.6	98.8	100.0	0.0
202	2061	55	19	1	0	98.9	80.0	80.3	0.5	99.8	73.7	100.0	0.0
210	2423	22	195	10	0	98.0	81.8	28.6	1.8	99.4	92.3	100.0	0.0
212	2748	0	0	0	0	99.5	–	0.0	0.5	100.0	–	–	0.0
213	2641	28	220	362	0	97.3	78.6	22.7	2.5	99.7	95.5	100.0	0.0
214	2002	0	256	1	2	99.8	–	0.0	0.2	99.5	95.5	100.0	0.0
219	2082	7	64	1	0	93.9	85.7	4.5	6.0	99.4	81.2	100.0	0.0
221	2031	0	396	0	0	97.8	–	0.0	2.2	99.3	96.0	100.0	0.0
222	2274	209	0	0	0	93.7	59.8	61.6	3.3	100.0	–	–	0.0
228	1688	3	362	0	0	94.6	100.0	3.2	5.4	99.3	96.1	100.0	0.0
231	1568	1	2	0	0	98.0	100.0	3.1	2.0	99.9	50.0	100.0	0.0
232	398	1382	0	0	0	87.3	81.3	100.0	0.0	100.0	–	–	0.0
233	2230	7	831	11	0	97.2	98.4	48.4	2.8	99.0	96.4	100.0	0.0
234	2700	50	3	0	0	98.5	84.0	56.0	1.2	100.0	66.7	100.0	0.0
sum	44258	1837	3221	388	7								
average						97.4	84.5	24.4	1.8	99.8	87.7	100.0	0.0

results for the VEB and SVEB classes are very promising indicating improvements over similar published studies.

Note that the AAMI recommended measures focus the attention on the classification of ventricular (VEB) and supraventricular heartbeats (SVEB). Table IV shows the results obtained using the AAMI guidelines for grouping classes in a multi-class scenario and the datasets stated in Table II. In order to compute performance measures for each class, we separate the multi-class confusion matrix into binary confusion matrices. Then measures such as sensitivity, specificity, accuracy and positive predicted value are computed using the typical definitions, see, for instance, [42] for the definition of these measures.

Table IV shows the performance of our system by class. In our approach all measures are higher than 84%. Sensitivity is particularly significant for clinical use since it counts for the percentage of true positive samples that were classified correctly. It is worth to remember that for this database no preprocessing or feature extraction were made over the signal, so a raw noisy waveform is being used for training and classifying. Table IV also shows the average performance reported by De Chazal et al. [9] and Llamedo et al. [10]. Reported performance was computed using their confusion matrices and separating them into binary matrices, to be consistent with our way of computing performance. De Chazal et al. [9] used feature extraction and linear discriminants (LD) to construct an automated heartbeat classifier. Although several configurations of the database were considered, we have compared our results with those of the best performance classifier. Unlike De Chazal et al., Llamedo et al. [10] utilized feature selection and signal modeling in order to build their classifiers. They used several databases, including the arrhythmia MIT-BIH database, however modifying the AAMI guidelines. They discarded AAMI class Q arguing that it is marginally represented in the database. They also merged AAMI classes F and V into a ventricular class (V'). Llamedo's work is the first attempt to classify heartbeats across databases.

TABLE IV
RESULTS FOR THE MULTICLASS CLASSIFICATION PROBLEM USING THE AAMI GUIDELINES FOR CLASS GROUPING. COMPARISON WITH OTHER METHODS THAT USED DS2.

Class	Sensitivity (%)	Specificity (%)	Accuracy (%)	$+P$ (%)
N	96.82	91.89	96.28	98.98
S	79.37	96.93	96.28	49.80
V	96.06	99.97	99.71	99.49
F	92.26	99.97	99.91	95.47
Q	57.14	100.00	99.99	100.00
average				
This article	**84.83**	**97.75**	**98.43**	**88.75**
LD [9]	65.95	96.04	94.35	45.57
LD/SM [10]	83.33	93.67	89.00	58.25

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2014.2332001, IEEE Journal of Biomedical and Health Informatics

TABLE V
CONFUSION MATRIX FOR THE AAMI-CLASSES CLASSIFICATION PROBLEM USING DS2

		Predicted samples				
		N	S	V	F	Q
Known samples	N	42852	1406	0	0	0
	S	379	1458	0	0	0
	V	53	60	3094	14	0
	F	10	4	16	358	0
	Q	0	0	0	3	4

Table V shows the multi-class confusion matrix of DS2 using the 5 AAMI classes. This table provides information on the misclassification of the different classes and serves as the base for future comparisons.

V. CONCLUSION

In this work, a state-of-the-art method for processing information known as Reservoir Computing has been adapted to study the open problem of the classification of heartbeats. Our approach requires a computationally inexpensive preprocessing of the data and is based only on the morphology of the heartbeat. Following standard recommendations, heartbeat types were grouped into five classes and several performance quantities were computed. Despite the different criteria mentioned in this work to facilitate comparisons with other studies, we consider that using the AAMI guidelines and balanced datasets for training and testing are essential for comparing algorithm performance among research teams. Our multi-class classification results indicate an average specificity of 97.75% with an average accuracy of 98.43%. Sensitivity and positive preditivity reach an average of 84.83% and 88.75%, respectively. In no case pre-processing or feature extraction step were implemented. However, a generalization of this approach to other databases will require the detection of an estimated position of the R-peak. Our results highlight the potential of the Resevoir Computing technique in the heartbeat classification problem. The results reported in this manuscript have been obtained using a single lead. The analysis of additional leads could provide a better understanding of the relevance of each lead for a proper heartbeat classification. This analysis, however, lies beyond the scope of this manuscript.

REFERENCES

[1] H. P. Selker, R. J. Zalenski, E. M. Antman, T. P. Aufderheide, S. A. Bernard, R. O. Bonow, W. B. Gibler, M. D. Hagen, P. Johnson, J. Lau *et al.*, "An evaluation of technologies for identifying acute cardiac ischemia in the emergency department: executive summary of a national heart attack alert program working group report," *Annals of Emergency Medicine*, vol. 29, no. 1, pp. 1–12, 1997.

[2] J. H. Pope and H. P. Selker, "Diagnosis of acute cardiac ischemia," *Emergency medicine clinics of North America*, vol. 21, no. 1, pp. 27–59, 2003.

[3] P. Brugada and J. Brugada, "Right bundle branch block, persistent st segment elevation and sudden cardiac death: a distinct clinical and electrocardiographic syndrome: a multicenter report," *Journal of the American College of Cardiology*, vol. 20, no. 6, pp. 1391–1396, 1992.

[4] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *Biomedical Engineering, IEEE Transactions on*, vol. BME-32, no. 3, pp. 230–236, 1985.

[5] Y. H. Hu, W. J. Tompkins, J. L. Urrusti, V. X. Afonso *et al.*, "Applications of artificial neural networks for ecg signal detection and classification," *Journal of electrocardiology*, vol. 26, pp. 66–73, 1993.

[6] S. Osowski and T. H. Linh, "Ecg beat recognition using fuzzy hybrid neural network," *Biomedical Engineering, IEEE Transactions on*, vol. 48, no. 11, pp. 1265–1271, 2001.

[7] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sornmo, "Clustering ecg complexes using hermite functions and self-organizing maps," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 7, pp. 838–848, 2000.

[8] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ecg beat classifier using a mixture of experts approach," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 9, pp. 891–900, 1997.

[9] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 7, pp. 1196–1206, 2004.

[10] M. Llamedo and J. Martinez, "Heartbeat classification using feature selection driven by database generalization criteria," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 3, pp. 616–625, 2011.

[11] C. Ye, M. T. Coimbra, and B. V. Kumar, "Investigation of human identification using two-lead electrocardiogram (ecg) signals," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*. IEEE, 2010, pp. 1–8.

[12] Z. Zidelmal, A. Amirou, D. Ould-Abdeslam, and J. Merckle, "Ecg beat classification using a cost sensitive classifier," *Computer methods and programs in biomedicine*, 2013.

[13] M. Engin, "Ecg beat classification using neuro-fuzzy network," *Pattern Recognition Letters*, vol. 25, no. 15, pp. 1715–1722, 2004.

[14] G. Karraz and G. Magenes, "Automatic classification of heartbeats using neural network classifier based on a bayesian framework," in *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, 2006, pp. 4016–4019.

[15] A. K. Mishra and S. Raghav, "Local fractal dimension based ecg arrhythmia classification," *Biomedical Signal Processing and Control*, vol. 5, pp. 114–123, 2010.

[16] C. Lainscsek and T. J. Sejnowski, "Electrocardiogram classification using delay differential equations," *Chaos*, vol. 23, no. 2, pp. 023 132–023 132–9, 2013.

[17] R. Mark and R. Wallen, "Aami-recommended practice: Testing and reporting performance results of ventricular arrhythmia detection algorithms," *Association for the Advancement of Medical Instrumentation, Arrythmia Monitorring Subcommittee, AAMI ECAR*, 1987.

[18] "Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms," *Association for the Advancement of Medical Instrumentation*, 1998.

[19] H. Jaeger, "The" echo state" approach to analysing and training recurrent neural networks-with an erratum note'," *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, vol. 148, 2001.

[20] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural computation*, vol. 14, no. 11, pp. 2531–2560, 2002.

[21] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, 2004.

[22] D. Verstraeten, B. Schrauwen, M. dHaene, and D. Stroobandt, "An experimental unification of reservoir computing methods," *Neural Networks*, vol. 20, no. 3, pp. 391–403, 2007.

[23] B. D. V. and M. W., "Statedependent computations: spatiotemporal processing in cortical network," *Nature Review Neurosciences*, vol. 10, pp. 113–125, 2009.

[24] W. Maass, P. Joshi, and E. D. Sontag, "Computational aspects of feedback in neural circuits," *PLOS Computational Biology*, vol. 3, no. 1, p. e165, 2007.

[25] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nature communications*, vol. 2, p. 468, 2011.

[26] M. Rabinovich, R. Huerta, and G. Laurent, "Transient dynamics for neural processing," *Science*, vol. 321, no. 5885, pp. 48–50, 2008.

[27] K. Hicke, M. Escalona-Morán, D. Brunner, M. C. Soriano, I. Fischer, and C. Mirasso, "Information processing using transient dynamics of semiconductor lasers subject to delayed feedback," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 4, p. 1501610, July-Aug 2013.

[28] E. A. Antonelo, B. Schrauwen, and D. Stroobandt, "Event detection and localization for small mobile robots using reservoir computing," *Neural Networks*, vol. 21, no. 6, pp. 862–871, 2008.
[29] A. A. Ferreira, T. B. Ludermir, R. R. de Aquino, M. M. Lira, and O. N. Neto, "Investigating the use of reservoir computing for forecasting the hourly wind speed in short-term," in *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*. IEEE, 2008, pp. 1649–1656.
[30] F. Wyffels and B. Schrauwen, "A comparative study of reservoir computing strategies for monthly time series prediction," *Neurocomputing*, vol. 73, no. 10, pp. 1958–1964, 2010.
[31] P. Coulibaly, "Reservoir computing approach to great lakes water level forecasting," *Journal of Hydrology*, vol. 381, no. 1, pp. 76–88, 2010.
[32] R. J. Abrahart, N. J. Mount, and A. Y. Shamseldin, "Discussion of reservoir computing approach to great lakes water level forecastingby p. coulibaly [j. hydrol. 381 (2010) 76–88]," *Journal of Hydrology*, vol. 422, pp. 76–80, 2012.
[33] P. Butenners, D. Verstraeten, B. Van Nieuwenhuyse, D. Stroobandt, R. Raedt, K. Vonck, P. Boon, and B. Schrauwen, "Real-time detection of epileptic seizures in animal models using reservoir computing," *Epilepsy Research*, vol. 103, pp. 124–134, 2013.
[34] E. T., *Applied Delayed Differential Equations*. Springer Science and Business Media, 2009.
[35] I. K. and M. K., "High-dimensional chaotic behavior in systems with time-delayed feedback," *Physica*, vol. D29, pp. 223–235, 1987.
[36] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nature communications*, vol. 4, p. 1364, January 2013.
[37] A. Rodan and P. Tino, "Minimum complexity echo state network," *Neural Networks, IEEE Transactions on*, vol. 22, no. 1, pp. 131–144, 2011.
[38] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. Wiley. com, 2013.
[39] Y. Pawitan, *In all likelihood: statistical modelling and inference using likelihood*. Oxford University Press, 2001.
[40] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, no. 4300, pp. 287–289, 1977.
[41] R. Mark, P. Schluter, G. Moody, P. Devlin, and D. Chernoff, "An annotated ecg database for evaluating arrhythmia detectors," *Front. Eng. Health Care*, pp. 205–210, 1982.
[42] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.

Miguel Angel Escalona-Morán received a BSc and MSc degrees in Physics from Universidad de Los Andes, Mérida - Venezuela, in 2004 and 2006 respectively.

In 2009, he won a Professorship in Mathematics at the Universidad de Los Andes. Currently Miguel Angel is pursuing a Ph.D. at the Institute for Cross-Disciplinary Physics and Complex Systems, joint center of the Spanish National Research Council and the University of the Balearic Islands, Palma de Mallorca, Spain. M. Escalona Morán has worked in collaboration with leading institutions such as the Central Bank of Venezuela and the Massachusetts Institute of Technology. His research interests cover fields from neuroscience to chaos and complex systems, including also nonlinear dynamics, biomedical engineering, time series analysis, modeling and numerical simulations.

Miguel Cornelles Soriano was born in Benicarlo, Spain, in 1979. He is member of the IEEE Photonics Society. He received the Telecommunications Engineering degree from the Universitat Politecnica de Catalunya, Barcelona, Spain, in 2002 and the Ph.D. degree in applied sciences from the Vrije Universiteit Brussel, Brussels, Belgium, in 2006.

He holds an assistant professor position at the Instituto de Física Interdisciplinar y Sistemas Complejos, Palma de Mallorca, Spain, since January 2011, where he previously held a 'Juan de la Cierva' scientific contract. His main research interests include the experimental and theoretical study of semiconductor lasers subject to delayed optical feedback and the synchronization properties of delay-coupled semiconductor lasers in the chaotic regime. Additional research lines cover topics such as information processing based on reservoir computing and nonlinear dynamics. As an author or co-author, he has published over 30 research papers in international refereed journals.

Ingo Fischer received the Diploma and Ph.D. degrees in physics from Philipps University, Marburg, Germany, in 1992 and 1995, respectively.

He stayed at Technical University of Darmstadt, Darmstadt, Germany, from 1995 to 2004, and at the Vrije Universiteit Brussel, Brussels, Belgium, from 2005 to 2007. In 2007, he became a Full Professor in photonics at Heriot-Watt University, Edinburgh, U.K. Since 2009, he has been a Professor at the Institute for Cross-Disciplinary Physics and Complex Systems, joint center of the Spanish National Research Council and the University of the Balearic Islands, Palma de Mallorca, Spain. His current research interests include nonlinear photonics and bio-inspired information processing, and in particular, the emission properties and dynamics of modern photonic sources, coupled laser systems, synchronization of lasers and neurons, and utilization of chaos.

Prof. Fischer received the Research Prize of the Adolf-Messer Foundation in 2000, and the first Hessian Cooperation Prize of the Technology Transfer Network in 2004.

Claudio R. Mirasso received the Ph.D. in physics from the Universidad Nacional de La Plata, Argentina, in 1989.

He has held post-doctoral positions in Spain and the Netherlands. He is Full Professor at the Physics Department, Universitat de les Illes Balears, Palma de Mallorca, Spain and researcher of the Institute for Cross-Disciplinary Physics and Complex Systems, joint center of the Spanish National Research Council and the University of the Balearic Islands. He has authored or co-authored over 140 journal papers. He was coordinator of the European Projects OCCULT and PHOCUS. His current research interests include dynamics of semiconductor lasers, synchronization and control of dynamical systems, dynamics and applications of delayed coupled systems, information processing, neuronal dynamics and applications of nonlinear dynamics in general.